

# The Cauchy-Schwarz Divergence for Assessing Situational Information Gain

Kyle J. DeMars  
NRC Postdoctoral Fellow  
Air Force Research Laboratory  
Kirtland AFB, NM 87111  
Email: demars.kyle@ieee.org

Islam I. Hussein  
Worcester Polytechnic Institute  
Department of Mechanical Engineering  
Worcester, MA 01609  
Email: ihussein@wpi.edu

Moriba K. Jah, R. Scott Erwin  
Air Force Research Laboratory  
Kirtland AFB, NM 87117

**Abstract**—In this paper, we consider the evaluation of information divergence and information gain as they apply to a hybrid random variable (i.e. a random variable which has both discrete and continuous elements) for multi-target tracking problems. In particular, we develop a closed-form solution for the Cauchy-Schwarz information divergence under the assumption that the continuous element of the random variable may be represented by a Gaussian mixture distribution and present the associated relationships for evaluating the Cauchy-Schwarz information gain. The developed information gain relationships are applied to a 0-1 target tracking problem common to space object tracking to determine the sensitivities to the information gain due to probability of detection, prior probability of object existence, and measurement noise.

## I. INTRODUCTION

One of the core concerns in Space Situational Awareness (SSA) is the maintenance of a catalog of tracked objects. Since the first launch of artificial satellites, the number of objects in orbit coming from new launches, decommissioned satellites, and debris created by collision of objects in orbit has posed an ever increasing challenge to the development of space object catalogs. As of 2006, there were approximately 9000 space objects being tracked by the U.S. Space Surveillance Network and maintained in the satellite catalog [1]. Currently, there are approximately 20,000 space objects currently being tracked, with 1000 of those objects being active objects. Furthermore, it is estimated that 500,000 objects with a diameter larger than one centimeter are in orbit. These numbers will inevitably increase as more objects are launched and as more collisions occur. The current number of objects coupled with the rapid advances in sensor technology that enable the detection of larger numbers of objects leads to a need for advanced strategies for scheduling sensors so as to optimally utilize available resources while maintaining accurate catalogs of space objects.

The current measure of performance for tasking is based upon maximizing the number of observations per prioritized objects. Given the scarcity of sensing resources, this metric will fail to consistently acquire objects for a growing number of detections. Mitigating this situation requires a method for dynamically assigning which targets are to be tracked and when they are to be tracked by a subset of the available sensors. The process of dynamic sensor tasking typically employs some measure of the information content of each available

sensor-object pair in order to formulate an optimization problem which schedules the sensors in a manner that maximizes the information gained regarding any individual object. In these problems, the actual measurements may be providing different types and qualities of data (e.g. line-of-sight data or range data). Additionally, since the sensors in a given network are neither identical or collocated their object information content is dependent on the dynamic environment, the sensor's location, orientation, and inherent accuracy. Therefore, the amount of information that can be gained on an object is not only a function of the target, but also of the sensor, and of the overall problem geometry.

Previous studies have examined the utilization of myopic algorithms for dynamic sensor tasking. For example, Erwin, *et al.* detailed the implementation of Fisher information as a measure of the information content in orbit determination problems [2]. Subsequently, Williams, *et al.* extended this approach to incorporate the utilization of the largest Lyapunov exponent in orbit determination problems [3]. Kreucher, *et al.* examined the general problem of information based sensor management from an information-theoretic perspective utilizing the Kullback-Leibler and Rényi divergences to formulate measures of information gain [4]. Extending the work of Kreucher, *et al.*, but in the context of sensor scheduling for antisubmarine warfare, Aughenbaugh and La Cour utilized information-theoretic information gain relationships for the Kullback-Leibler and Rényi divergences to assess the performance of myopic sensor scheduling problems [5]. Extending the work of Aughenbaugh and La Cour, DeMars and Jah developed and investigated the utilization of information gain measures for several class of information-theoretic divergences for the problem of sensor tasking in uncertain orbital dynamical systems [6].

It is important to note that the utilization of sensor time is both scarce and expensive. Decisions on whether to operate a sensor in a mode which optimizes tracking capabilities versus a mode which optimizes detection capabilities require an assessment of how much information can be extracted (or gained). This expected information gain will have two (interdependent) components: one that is continuous in nature and another that is discrete in nature. Therefore, the expected information gain associated with each sensor assignment is

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE <b>JUL 2012</b>	2. REPORT TYPE		3. DATES COVERED <b>00-00-2012 to 00-00-2012</b>		
4. TITLE AND SUBTITLE <b>The Cauchy-Schwarz Divergence for Assessing Situational Information Gain</b>			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Air Force Research Laboratory, Kirtland AFB, NM, 87117</b>			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES <b>Presented at the 15th International Conference on Information Fusion held in Singapore on 9-12 July 2012. Sponsored in part by Office of Naval Research and Office of Naval Research Global.</b>					
14. ABSTRACT <b>In this paper, we consider the evaluation of information divergence and information gain as they apply to a hybrid random variable (i.e. a random variable which has both discrete and continuous elements) for multi-target tracking problems. In particular, we develop a closed-form solution for the Cauchy-Schwarz information divergence under the assumption that the continuous element of the random variable may be represented by a Gaussian mixture distribution and present the associated relationships for evaluating the Cauchy-Schwarz information gain. The developed information gain relationships are applied to a 0-1 target tracking problem common to space object tracking to determine the sensitivities to the information gain due to probability of detection, prior probability of object existence and measurement noise.</b>					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>Same as Report (SAR)</b>	18. NUMBER OF PAGES <b>8</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

of a hybrid nature. The final requirement, therefore, for an effective approach to solving the SSA problem must provide rigorous machinery for quantifying hybrid information gain for optimal sensor allocation.

The goal of this paper is to investigate the Cauchy-Schwarz information divergence and its associated information gain within the multi-target tracking framework of Finite Set Statistics (FISST) [7], [8]. Specifically, we develop a closed-form solution for the Cauchy-Schwarz information divergence, present a method for determining the associated information gain, and apply the developments to a multi-target tracking problem.

The paper is organized as follows: the problem statement and relevant notation is provided in Section II, the basic formulation of the Cauchy-Schwarz divergence for multi-target problems and a closed-form solution are given in Section III, a discussion of the Cauchy-Schwarz information gain is given in Section IV, some results are presented in Section V, and we conclude with some remarks in Section VI.

## II. PROBLEM STATEMENT

As opposed to purely discrete or purely continuous Bayesian inference, FISST makes use of set-valued random variables. An example of a set-valued random variable is the state  $X = \{x_d, \mathbf{x}\}$  in an SSA characterization and tracking inference problem. If we let  $W$  be the set of all possible object types, then  $x_d \in W$  is the discrete component of the state that describes a space object's type (and, hence, its dynamic model) and  $\mathbf{x} \in \mathbb{R}^s$  is the continuous component of the  $s$ -dimensional state (e.g. the position and velocity of an object). In detection and tracking, the system state  $X = (n, \mathbf{X})$ , where  $n$  is the discrete component of the state that describes the number of objects in the search space and  $\mathbf{X}^T = [\mathbf{x}_1^T \ \mathbf{x}_2^T \ \dots \ \mathbf{x}_n^T] \in \mathbb{R}^{sn}$  describes the positions and velocities of these objects. Notice here the explicit dependence of the dimension of the continuous state space  $\mathbb{R}^{sn}$  on the discrete component  $n$  of the state. For brevity, we simply write  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ .

Bayes' law for performing a measurement update step takes on exactly the same form in the hybrid FISST approach as it does in purely continuous and or purely discrete problems, that is

$$f_{k+1|k+1}(X|Z^{(k+1)}) \propto f_{k+1}(Z_{k+1}|X)f_{k+1|k}(X|Z^{(k)}) \quad (1)$$

where  $f_{k+1}(Z|X)$  is the multi-target likelihood function that describes the likelihood of getting a measurement  $Z_{k+1}$  given the state  $X_{k+1}$ , and  $Z^{(k)} : Z_1, \dots, Z_k$  is the time sequence of measurement sets up to time  $k$ . If desired, Eq. (1) can be changed to an equality by dividing the right-hand side by the Bayes' factor, which is given by

$$f_{k+1}(Z_{k+1}|Z^{(k)}) = \int f_{k+1}(Z_{k+1}|X)f_{k+1|k}(X|Z^{(k)})\delta X \quad (2)$$

Notice that the integrals are set integrals. For multi-target detection and tracking, a set integral of a scalar-valued set

function  $g(X)$  is defined to be the integral of  $g$  over the continuous component, summed over all possible discrete values [7], [8]

$$\int g(X)\delta X = g(X = \emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int g(\{\mathbf{x}_1, \dots, \mathbf{x}_n\})d\mathbf{x}_1 \dots d\mathbf{x}_n \quad (3)$$

where the factorial coefficient is to take into account all the different possible orderings of  $X$  as evaluated in the function  $g$ .

In order to develop measures of the information gain available by scheduling measurements, we first consider measures of the directed difference between two generalized pdfs, namely the *a priori* and *a posteriori* pdfs, i.e. the generalized pdfs immediately before and after processing measurement data, respectively. Generally speaking, the information divergence is a measure of distance (i.e. similarity or dissimilarity) between two pdfs. Given an information divergence describing the directed distance from  $p(X)$  to  $q(X)$  denoted by  $D[p||q]$ , the "distance" is called a metric if [9]

- 1)  $D[p||q] \geq 0$  with equality iff  $p(X) = q(X)$  (non-negativity and positive definiteness),
- 2)  $D[p||q] = D[q||p]$  (symmetry), and
- 3)  $D[p||r] \leq D[p||q] + D[q||r]$  (sub-additivity/triangle inequality).

Information divergences which only satisfy the first condition are referred to as asymmetric divergences, whereas satisfaction of the second condition necessarily removes the restriction of referring to the divergence as asymmetric. However, in this work, asymmetric divergences will be referred to as divergences with the understanding that symmetry is not required for the results to hold. One of the most common information divergences is the Kullback-Leibler divergence, given by [10]

$$D_{KL}[p||q] = \int p(X) \log \frac{p(X)}{q(X)} \delta X \quad (4)$$

which was investigated for multi-target tracking by Uney *et al.* [11]. The Kullback-Leibler divergence, however, only admits closed-form solutions in special cases, such as for linear Gaussian systems. Motivated by this fact, we consider the Cauchy-Schwarz divergence which has a closed-form solution for single target tracking frameworks [6].

## III. CAUCHY-SCHWARZ INFORMATION DIVERGENCE

By defining a the inner-product of two square-integrable functions  $p(X)$  and  $q(X)$  as  $\langle p, q \rangle = \int p(X)q(X)\delta X$ , the Cauchy-Schwarz inequality may be used to define the Cauchy-Schwarz information divergence as [12]

$$\begin{aligned} D_{CS}[p||q] &= \frac{1}{2} \log \left[ \frac{(\int p^2(X)\delta X)(\int q^2(X)\delta X)}{(\int p(X)q(X)\delta X)^2} \right] \\ &= \frac{1}{2} \log \int p^2(X)\delta X + \frac{1}{2} \log \int q^2(X)\delta X \\ &\quad - \log \int p(X)q(X)\delta X \end{aligned} \quad (5)$$

where, for the purposes of this work,  $q(X)$  represents the multi-target prior generalized pdf and  $p(X)$  represents the multi-target posterior generalized pdf. The Cauchy-Schwarz divergence may also be generalized to a class of information divergences via introduction of a control parameter; this class of information divergences is known as the Gamma divergence, for which the Cauchy-Schwarz divergence is a special case [9]. Additionally, the Cauchy-Schwarz divergence is implicitly related to the quadratic entropy of Rényi, which, for pdf  $r(X)$  is given by [13]

$$H_R^{(2)} = -\log \int r^2(X) \delta X$$

which is of the same form as the first two quantities in Eq. (5). The Cauchy-Schwarz does not satisfy the triangle inequality and therefore cannot be classified as a metric, but it does satisfy the following properties [14]:

- 1)  $D_{CS}[p||q] \geq 0 \forall p, q$
- 2)  $D_{CS}[p||q] = 0$  iff  $f(X) = g(X)$
- 3)  $D_{CS}[p||q] = D_{CS}[q||p]$
- 4)  $D_{CS}[p||q]$  is additive for independent events
- 5)  $D_{CS}[c p||q] = D_{CS}[p||q]$  for any  $c > 0$

These properties illustrate that the Cauchy-Schwarz divergence is positive semi-definite, symmetric, and scale-invariant. This last property is a very nice feature of the Cauchy-Schwarz divergence which we will make use of in the sequel.

For the sake of brevity and ease of notation, we restrict our attention to a case in which there can exist at most one object and at most one clutter point in the search space, which we refer to as the “0-1 problem”. A summary of the pertinent FISST equations is provided in the Appendix, and a full treatment of the development of the FISST equations for the 0-1 problem is given in Hussein, *et al.* [15].

Since we are considering the 0-1 problem, we need only to account for the possibilities that there is no target, i.e.  $X = \emptyset$ , and that there is a single target, i.e.  $X = \{x\}$ . Then, by Eq. (3), the integral terms of Eq. (5) may be written as

$$\begin{aligned} \int p^2(X) \delta X &= p^2(\emptyset) + \int p^2(x) dx \\ \int q^2(X) \delta X &= q^2(\emptyset) + \int q^2(x) dx \\ \int p(X)q(X) \delta X &= p(\emptyset)q(\emptyset) + \int p(x)q(x) dx \end{aligned}$$

At this point, the prior generalized pdf is associated with  $q(X)$  and the posterior generalized pdf is associated with  $p(X)$ , such that

$$\begin{aligned} p(\emptyset) &= f_{k+1|k+1}(X = \emptyset|Z^{(k+1)}) \\ p(x) &= f_{k+1|k+1}(X = \{x\}|Z^{(k+1)}) \\ q(\emptyset) &= f_{k+1|k}(X = \emptyset|Z^{(k)}) \\ q(x) &= f_{k+1|k}(X = \{x\}|Z^{(k)}) \end{aligned}$$

which allows the Cauchy-Schwarz divergence for the 0-1 problem to be expressed as

$$\begin{aligned} D_{CS} &= \frac{1}{2} \log \left[ f_{k+1|k+1}^2(X = \emptyset|Z^{(k+1)}) \right. \\ &\quad + \int f_{k+1|k+1}^2(X = \{x\}|Z^{(k+1)}) dx \\ &\quad + \frac{1}{2} \log \left[ f_{k+1|k}^2(X = \emptyset|Z^{(k)}) \right. \\ &\quad + \int f_{k+1|k}^2(X = \{x\}|Z^{(k)}) dx \\ &\quad - \log \left[ f_{k+1|k+1}(X = \emptyset|Z^{(k+1)}) \right. \\ &\quad \times f_{k+1|k}(X = \emptyset|Z^{(k)}) \\ &\quad + \int f_{k+1|k+1}(X = \{x\}|Z^{(k+1)}) \\ &\quad \times f_{k+1|k}(X = \{x\}|Z^{(k)}) dx \left. \right] \end{aligned} \quad (6)$$

Note that in Eq. (6) we have dropped the functional dependence of the Cauchy-Schwarz divergence on the pdfs for which the divergence is computed as it is no longer ambiguous which pdfs are the inputs. Recalling the scale-invariance property of the Cauchy-Schwarz divergence and substituting for  $f_{k+1|k+1}(X|Z^{(k+1)})$  in Eq. (6) from the Bayes’ rule update of Eq. (1) yields

$$\begin{aligned} D_{CS} &= \frac{1}{2} \log \left[ f_{k+1}^2(Z_{k+1}|X = \emptyset) f_{k+1|k}^2(X = \emptyset|Z^{(k)}) \right. \\ &\quad + \int f_{k+1}^2(Z_{k+1}|X = \{x\}) f_{k+1|k}^2(X = \{x\}|Z^{(k)}) dx \\ &\quad + \frac{1}{2} \log \left[ f_{k+1|k}^2(X = \emptyset|Z^{(k)}) \right. \\ &\quad + \int f_{k+1|k}^2(X = \{x\}|Z^{(k)}) dx \\ &\quad - \log \left[ f_{k+1}(Z_{k+1}|X = \emptyset) f_{k+1|k}(X = \emptyset|Z^{(k)}) \right. \\ &\quad + \left. \int f_{k+1}(Z_{k+1}|X = \{x\}) f_{k+1|k}(X = \{x\}|Z^{(k)}) dx \right] \end{aligned}$$

Now, we must consider different measurement sets independently. Since we have restricted our attention to the 0-1 problem, three possible measurement sets are possible:

- 1) no sensor return, in which case  $Z_{k+1} = \emptyset$
- 2) a single sensor return, in which case  $Z_{k+1} = \{z\}$
- 3) two sensor returns, in which case  $Z_{k+1} = \{z_1, z_2\}$

For each case, we apply the 0-1 problem FISST equations that are summarized in the Appendix and developed by Hussein, *et al.* [15]. Before proceeding, it is useful to define some terms which appear repeatedly. Let  $p$  be the prior probability that the object exists at time  $k$ ,  $p_D$  be the probability of detection, and  $p_F$  be the probability of false alarm. Furthermore, let us

define  $\mathcal{I}_0$ ,  $\mathcal{I}_2(\mathbf{z})$ , and  $\mathcal{I}_2(\mathbf{z}_1, \mathbf{z}_2)$  as

$$\mathcal{I}_0 = \int f_{k+1|k}^2(\mathbf{x}|Z^{(k)})d\mathbf{x} \quad (7)$$

$$\mathcal{I}_1(\mathbf{z}) = \int f_{k+1}(\mathbf{z}|\mathbf{x})f_{k+1|k}^2(\mathbf{x}|Z^{(k)})d\mathbf{x} \quad (8)$$

$$\mathcal{I}_2(\mathbf{z}_1, \mathbf{z}_2) = \int f_{k+1}(\mathbf{z}_1|\mathbf{x})f_{k+1}(\mathbf{z}_2|\mathbf{x})f_{k+1|k}^2(\mathbf{x}|Z^{(k)})d\mathbf{x} \quad (9)$$

For the case of no sensor return ( $Z_{k+1} = \emptyset$ ), it can be shown that

$$\begin{aligned} D_{CS}(Z_{k+1} = \emptyset) &= \frac{1}{2} \log \left[ (1-p)^2 + p^2(1-p_D)^2 \mathcal{I}_0 \right] \\ &\quad + \frac{1}{2} \log \left[ (1-p)^2 + p^2 \mathcal{I}_0 \right] \\ &\quad - \log \left[ (1-p)^2 + p^2(1-p_D) \mathcal{I}_0 \right] \end{aligned} \quad (10)$$

For the case of a single sensor return ( $Z_{k+1} = \{\mathbf{z}\}$ ), it can be shown that

$$\begin{aligned} D_{CS}(Z_{k+1} = \{\mathbf{z}\}) &= \\ &\frac{1}{2} \log \left[ (1-p)^2 p_F^2 g^2(\mathbf{z}) + p^2 p_F^2 (1-p_D)^2 g^2(\mathbf{z}) \mathcal{I}_0 \right. \\ &\quad + 2p^2 p_F (1-p_F) p_D (1-p_D) g(\mathbf{z}) \mathcal{I}_1(\mathbf{z}) \\ &\quad \left. + p^2 (1-p_F)^2 p_D^2 \mathcal{I}_2(\mathbf{z}, \mathbf{z}) \right] + \frac{1}{2} \log \left[ (1-p)^2 + p^2 \mathcal{I}_0 \right] \\ &\quad - \log \left[ (1-p)^2 p_F g(\mathbf{z}) + p^2 p_F (1-p_D) g(\mathbf{z}) \mathcal{I}_0 \right. \\ &\quad \left. + p^2 (1-p_F) p_D \mathcal{I}_1(\mathbf{z}) \right] \end{aligned} \quad (11)$$

where  $g(\mathbf{z})$  is the spatial likelihood distribution function that a clutter point generated the measurement  $\mathbf{z}$ . For the case of two sensor returns ( $Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\}$ ), it can be shown that

$$\begin{aligned} D_{CS}(Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\}) &= \\ &\frac{1}{2} \log \left[ g^2(\mathbf{z}_1) \mathcal{I}_2(\mathbf{z}_2, \mathbf{z}_2) + g^2(\mathbf{z}_2) \mathcal{I}_2(\mathbf{z}_1, \mathbf{z}_1) \right. \\ &\quad \left. + 2g(\mathbf{z}_1)g(\mathbf{z}_2) \mathcal{I}_2(\mathbf{z}_1, \mathbf{z}_2) \right] + \frac{1}{2} \log \left[ (1-p)^2 + p^2 \mathcal{I}_0 \right] \\ &\quad - \log \left[ pg(\mathbf{z}_1) \mathcal{I}_1(\mathbf{z}_2) + pg(\mathbf{z}_2) \mathcal{I}_1(\mathbf{z}_1) \right] \end{aligned} \quad (12)$$

where  $g(\mathbf{z}_1)$  is the spatial likelihood distribution function that a clutter point generated the measurement  $\mathbf{z}_1$  and similarly for  $g(\mathbf{z}_2)$ .

Up to this point, no explicit forms of the pdfs involved in the computation of the Cauchy-Schwarz divergence have been introduced, rendering the preceding results completely general outside of the specification to the 0-1 problem. To obtain solutions which are readily implementable in computations, however, it is useful to specify forms of the involved pdfs so as to obtain a closed-form solutions for Eqs. (10)–(12). Specifically, this means that the forms of  $f_{k+1}(\mathbf{z}|\mathbf{x})$  and  $f_{k+1|k}(\mathbf{x}|Z^{(k)})$  need to be prescribed so that the integral terms of Eqs. (7)–(9) may be computed and utilized in Eqs. (10)–(12).

#### A. Closed-Form Solution of the Cauchy-Schwarz Divergence

To obtain closed-form solutions to the Cauchy-Schwarz divergence equations, we first assume that the prior pdf,  $f_{k+1|k}(\mathbf{x}|Z^{(k)})$ , and the measurement pdf,  $f_{k+1}(\mathbf{z}|\mathbf{x})$ , are represented by a Gaussian mixture and by a Gaussian, respectively, such that

$$f_{k+1|k}(\mathbf{x}|Z^{(k)}) = \sum_{i=1}^L w_i p_g(\mathbf{x}; \mathbf{m}_i, \mathbf{P}_i) \quad (13)$$

$$f_{k+1}(\mathbf{z}|\mathbf{x}) = p_g(\mathbf{z}; \mathbf{H}\mathbf{x}, \mathbf{R}) \quad (14)$$

where  $p_g(\mathbf{y}; \mathbf{a}, \mathbf{A})$  is used to denote a Gaussian pdf with mean  $\mathbf{a}$  and covariance  $\mathbf{A}$ . Before proceeding further, it is worth noting two identities regarding multiplying Gaussian pdfs. The product of two Gaussian pdfs of the same random variable is given by an unnormalized Gaussian pdf as [16]

$$p_g(\mathbf{x}; \mathbf{a}, \mathbf{A}) p_g(\mathbf{x}; \mathbf{b}, \mathbf{B}) = \Gamma(\mathbf{a}, \mathbf{b}, \mathbf{A}, \mathbf{B}) p_g(\mathbf{x}; \mathbf{c}, \mathbf{C}) \quad (15)$$

where

$$\begin{aligned} \mathbf{c} &= \mathbf{C}(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b}) \\ \mathbf{C} &= (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1} \\ \Gamma(\mathbf{a}, \mathbf{b}, \mathbf{A}, \mathbf{B}) &= |2\pi(\mathbf{A} + \mathbf{B})|^{-1/2} \\ &\quad \times \exp \left\{ -\frac{1}{2}(\mathbf{a} - \mathbf{b})^T (\mathbf{A} + \mathbf{B})^{-1} (\mathbf{a} - \mathbf{b}) \right\} \end{aligned}$$

The second identity states that for  $\mathbf{H}$ ,  $\mathbf{R}$ ,  $\mathbf{m}$ , and  $\mathbf{P}$  of matching dimensions with  $\mathbf{R}$  and  $\mathbf{P}$  positive definite [17]

$$p_g(\mathbf{z}; \mathbf{H}\mathbf{x}, \mathbf{R}) p_g(\mathbf{x}; \mathbf{m}, \mathbf{P}) = Q(\mathbf{z}; \mathbf{H}, \mathbf{m}, \mathbf{P}, \mathbf{R}) p_g(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (16)$$

where

$$\begin{aligned} \boldsymbol{\mu} &= \mathbf{m} + \mathbf{K}(\mathbf{z} - \mathbf{H}\mathbf{m}) \\ \boldsymbol{\Sigma} &= \mathbf{P} - \mathbf{K}\mathbf{H}\mathbf{P} \\ \mathbf{K} &= \mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1} \\ Q(\mathbf{z}; \mathbf{H}, \mathbf{m}, \mathbf{P}, \mathbf{R}) &= p_g(\mathbf{z}; \mathbf{H}\mathbf{m}, \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}) \end{aligned}$$

To obtain closed-form solutions to the Cauchy-Schwarz divergence of Eqs. (10)–(12), we only need to find closed-form solutions for the integral terms of Eqs. (7)–(9). We begin by noting that  $\mathcal{I}_0$  may be written as

$$\mathcal{I}_0 = \int f_{k+1|k}(\mathbf{x}|Z^{(k)}) f_{k+1|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$$

Then, substituting for  $f_{k+1|k}(\mathbf{x}|Z^{(k)})$  from Eq. (13) and applying the identity of Eq. (15), it follows that  $\mathcal{I}_0$  is given by

$$\mathcal{I}_0 = \sum_{i=1}^L \sum_{j=1}^L w_i w_j \Gamma(\mathbf{m}_i, \mathbf{m}_j, \mathbf{P}_i, \mathbf{P}_j) \quad (17)$$

In a similar approach to that of computing  $\mathcal{I}_0$ ,  $\mathcal{I}_1$  may be alternatively expressed as

$$\mathcal{I}_1(\mathbf{z}) = \int f_{k+1}(\mathbf{z}|\mathbf{x}) f_{k+1|k}(\mathbf{x}|Z^{(k)}) f_{k+1|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$$

Substituting for  $f_{k+1|k}(\mathbf{x}|Z^{(k)})$  from Eq. (13) and  $f_{k+1}(\mathbf{z}|\mathbf{x})$  from Eq. (14), then applying the identities of Eqs. (15) and (16), we obtain

$$\mathcal{I}_1(\mathbf{z}) = \sum_{i=1}^L \sum_{j=1}^L w_i w_j Q(\mathbf{z}; \mathbf{H}, \mathbf{m}_i, \mathbf{P}_i, \mathbf{R}) \quad (18)$$

$$\times \Gamma(\boldsymbol{\mu}_i, \mathbf{m}_j, \boldsymbol{\Sigma}_i, \mathbf{P}_j)$$

where

$$\begin{aligned} \boldsymbol{\mu}_i &= \mathbf{m}_i + \mathbf{K}_i(\mathbf{z} - \mathbf{H}\mathbf{m}_i) \\ \boldsymbol{\Sigma}_i &= \mathbf{P}_i - \mathbf{K}_i \mathbf{H} \mathbf{P}_i \\ \mathbf{K}_i &= \mathbf{P}_i \mathbf{H}^T (\mathbf{H} \mathbf{P}_i \mathbf{H}^T + \mathbf{R})^{-1} \end{aligned}$$

Finally,  $\mathcal{I}_2$  may be expressed as

$$\begin{aligned} \mathcal{I}_2(\mathbf{z}_1, \mathbf{z}_2) &= \int f_{k+1}(\mathbf{z}_1|\mathbf{x}) f_{k+1|k}(\mathbf{x}|Z^{(k)}) \\ &\quad \times f_{k+1}(\mathbf{z}_2|\mathbf{x}) f_{k+1|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x} \end{aligned}$$

Once again, by substituting for  $f_{k+1|k}(\mathbf{x}|Z^{(k)})$  from Eq. (13) and  $f_{k+1}(\mathbf{z}|\mathbf{x})$  from Eq. (14), then applying the identities of Eqs. (15) and (16), it can be shown that

$$\begin{aligned} \mathcal{I}_2(\mathbf{z}_1, \mathbf{z}_2) &= \sum_{i=1}^L \sum_{j=1}^L w_i w_j Q(\mathbf{z}_1, \mathbf{H}, \mathbf{m}_i, \mathbf{P}_i, \mathbf{R}) \quad (19) \\ &\quad \times Q(\mathbf{z}_2, \mathbf{H}, \mathbf{m}_j, \mathbf{P}_j, \mathbf{R}) \\ &\quad \times \Gamma(\boldsymbol{\mu}_{1,i}, \boldsymbol{\mu}_{2,j}, \boldsymbol{\Sigma}_i, \boldsymbol{\Sigma}_j) \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{1,i} &= \mathbf{m}_i + \mathbf{K}_i(\mathbf{z}_1 - \mathbf{H}\mathbf{m}_i) \\ \boldsymbol{\mu}_{2,i} &= \mathbf{m}_i + \mathbf{K}_i(\mathbf{z}_2 - \mathbf{H}\mathbf{m}_i) \\ \boldsymbol{\Sigma}_i &= \mathbf{P}_i - \mathbf{K}_i \mathbf{H} \mathbf{P}_i \\ \mathbf{K}_i &= \mathbf{P}_i \mathbf{H}^T (\mathbf{H} \mathbf{P}_i \mathbf{H}^T + \mathbf{R})^{-1} \end{aligned}$$

Thus, a closed-form solution to the Cauchy-Schwarz information divergence for the 0-1 problem has been obtained under the assumptions that the state pdf may be represented as a Gaussian mixture and that the measurement pdf may be represented as a Gaussian. To summarize, Eqs. (17), (18), and (19) are utilized to compute  $\mathcal{I}_0$ ,  $\mathcal{I}_2(\mathbf{z})$ , and  $\mathcal{I}_2(\mathbf{z}_1, \mathbf{z}_2)$ , which may then be employed in Eqs. (10)–(12) to compute the Cauchy-Schwarz information divergence, with the specific equation employed being dependent upon whether there were no sensor returns, a single sensor return, or two sensor returns.

#### IV. CAUCHY-SCHWARZ INFORMATION GAIN

The Cauchy-Schwarz information divergence provides a method by which the amount of acquired information regarding the state (both the discrete and continuous components) may be determined given measurement data (i.e. no return, a single return, or two returns). It does not, however, provide a measure that can be used to assess future performance, i.e. in the case that no data is yet available. For this reason, we define the Cauchy-Schwarz information gain to be the expected value of the information divergence over all possible measurement

outcomes. Since  $D_{CS} : X \times Z \mapsto \mathbb{R}^+$ , it is seen that by Eq. (3), the information gain may be written as

$$\begin{aligned} G_{CS} &= f(Z_{k+1} = \emptyset) D_{CS}(Z = \emptyset) \\ &\quad + \int f(Z_{k+1} = \{\mathbf{z}\}) D_{CS}(Z = \{\mathbf{z}\}) d\mathbf{z} \\ &\quad + \frac{1}{2} \int f(Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\}) D_{CS}(Z = \{\mathbf{z}_1, \mathbf{z}_2\}) d\mathbf{z}_1 d\mathbf{z}_2, \end{aligned}$$

where  $f(Z_{k+1} = \emptyset)$ ,  $f(Z_{k+1} = \{\mathbf{z}\})$ , and  $f(Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\})$  are the Bayes factors for the cases of no return, a single return, and two returns, respectively. The forms of the Bayes factors are given in the Appendix and discussed in more detail in Reference [15]. Notice here that we generalize the conventional definition of expectations to compute the expected hybrid information divergence. This is a mathematically well-defined operation since the information divergence function is a real-valued set-function. This generalization is mathematically ill-defined if the function one is taking an expectation of is set-valued, say in attempting to compute the expected value of a set-valued random variable  $X$ . For more on this, see Chapter 16 of Reference [8].

Letting  $p_0 = (1 - p_F)[(1 - p) + p(1 - p_D)]$ ,  $p_f = p p_D(1 - p_F)$ ,  $p_g = p_F[(1 - p) + p(1 - p_D)]$ , and  $p_{fg} = p p_F p_D$ , it follows that the Cauchy-Schwarz information gain may be expressed as

$$\begin{aligned} G_{CS} &= p_0 E_0 + p_f E_f(\mathbf{z}) + p_g E_g(\mathbf{z}) \quad (20) \\ &\quad + \frac{1}{2} p_{fg} E_{fg}(\mathbf{z}_1, \mathbf{z}_2) + \frac{1}{2} p_{fg} E_{fg}(\mathbf{z}_2, \mathbf{z}_1) \end{aligned}$$

where

$$\begin{aligned} E_0 &= D_{CS}(Z = \emptyset) \\ E_f(\mathbf{z}) &= \int f_{k+1}(\mathbf{z}) D_{CS}(Z = \{\mathbf{z}\}) d\mathbf{z} \quad (21) \end{aligned}$$

$$E_g(\mathbf{z}) = \int g(\mathbf{z}) D_{CS}(Z = \{\mathbf{z}\}) d\mathbf{z} \quad (22)$$

$$E_{fg}(\mathbf{z}_1, \mathbf{z}_2) = \int f_{k+1}(\mathbf{z}_1) g(\mathbf{z}_2) D_{CS}(Z = \{\mathbf{z}_1, \mathbf{z}_2\}) d\mathbf{z}_1 d\mathbf{z}_2 \quad (23)$$

and  $f_{k+1}(\mathbf{z})$  is the spatial likelihood distribution function that the target generated the measurement, which is given by

$$f_{k+1}(\mathbf{z}) = \int f_{k+1}(\mathbf{z}|\mathbf{x}) f_{k+1|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$$

Furthermore, it is reminded that  $g(\mathbf{z})$  is the spatial likelihood distribution function that a clutter point generated the measurement. As before, by substituting for  $f_{k+1|k}(\mathbf{x}|Z^{(k)})$  from Eq. (13) and  $f_{k+1}(\mathbf{z}|\mathbf{x})$  from Eq. (14), then applying the identities of Eqs. (15) and (16), it follows that  $f_{k+1}(\mathbf{z})$  can be written as

$$f_{k+1}(\mathbf{z}) = \sum_{i=1}^L w_i Q(\mathbf{z}; \mathbf{H}, \mathbf{m}_i, \mathbf{P}_i, \mathbf{R})$$

In general, the integral equations of Eqs. (21)–(23) admit no known closed-form solutions, and so we compute them via

monte carlo integration. Additionally, it should be noted that the information gain relationship in Eq. (20) naturally decomposes into contributions from no return which is represented in the first term, a single return (either target or clutter generated) through the second and third terms, and two returns in the fourth and fifth terms.

## V. RESULTS

Given the preceding results on computing both the information divergence and the associated information gain, there are several ways in which the methods may be applied. For instance, Reference [15] illustrates the information divergence as a function of time, illustrating the effectiveness of measurements in a multi-target tracking problem, and Reference [6] illustrates the information gain as a function of time as a potential mechanism for determining the times at which measurements can be taken to obtain maximum information gain.

In the sequel, we consider a fixed point in time (with a fixed continuous state pdf) and use the information gain as a method for determining the sensitivities to variations in the prior probability of object existence, the probability of detection, the probability of false alarm, and the measurement noise. Furthermore, to illustrate the flexibility of the developed methods, we apply the information gain calculations to two scenarios: 1) a target which is represented by a Gaussian distribution (in the continuous state) with a nearby object that can generate false alarms, also with a Gaussian distribution and 2) a target which is represent by a Gaussian mixture distribution (in the continuous state) with a uniform clutter distribution defined over a portion of the sensor field of view.

For both problems considered, the dynamical system model is that of a planar two-body orbital motion problem with a sensor that is on the surface of the Earth and can take measurements of the target's position. That is, the dynamical system is given by

$$\begin{bmatrix} \dot{\mathbf{r}} \\ \dot{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ -\mu \mathbf{r} r^{-3} \end{bmatrix},$$

where  $\mathbf{r}$  is the inertial position of the object,  $\mathbf{v}$  is the inertial velocity of the object, and  $\mu$  is the gravitational parameter of the Earth. Additionally, the measurements are taken to be of the form

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n},$$

where  $\mathbf{H}$  is such that  $\mathbf{H}\mathbf{x} = \mathbf{r}$ , and  $\mathbf{n}$  is the measurement noise, which is taken to be zero-mean with covariance  $\mathbf{R} = \sigma^2 \mathbf{I}_{2 \times 2}$ .

A schematic representing the observational geometry for the first scenario considered is given in Figure 1. The continuous state target pdf is characterized by a Gaussian distribution with 1 [km] position uncertainty and 1 [m/s] velocity uncertainty. Additionally, the mean is described by an apoapsis of 42,100 [km] and an eccentricity of 0.2. The clutter model in this case is represented by a nearby object that generates false returns with a pdf of  $g(\mathbf{z}) = p_g(\mathbf{z}, \mathbf{m}_c, \mathbf{R}_c)$ , where  $\mathbf{m}_c$

is chosen to be 100 [m] from the true object in both  $x$  and  $y$  positions and  $\mathbf{R}_c = (\sigma_c)^2 \mathbf{I}_{2 \times 2}$  with  $\sigma_c = 25$  [km]. The information gain as a function of the probability of detection is shown in Figure 2 for several values of the prior probability of object existence. This shows that for all values of  $p$ , an increase in  $p_D$  leads to higher information gain. Additionally, it is seen that for high values of  $p_D$  a larger information gain results from smaller  $p$ , which is largely due to the information gained on the probability of object existence. In Figure 3, the information gain is shown as a function of prior probability of object existence for several values of the measurement noise,  $\sigma$ . Here, it is seen that lower measurement noise leads to higher information gain across the range of  $p$ . Additionally, an interesting inflection point is observed for low values of  $p$ . To explain this effect, we show the contributions to the information gain in Figure 4, which correspond to each of the terms in Eq. (20). This shows that for  $p = 0$ , the information gain is zero, but for small non-zero values of  $p$ , the no return and clutter return contributions to the information gain are high. As  $p$  increases, these two contributions quickly decrease and the remaining contributions become dominant. The trade-off between the two trends causes the inflection observed in the information gain of Figure 3.

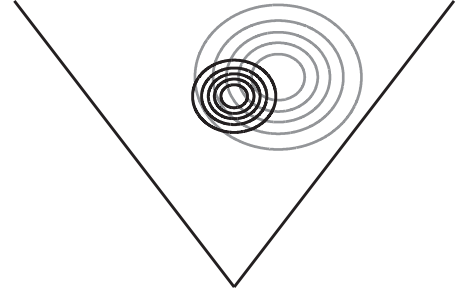


Fig. 1. Schematic of the Gaussian target/Gaussian clutter model. The black contour lines represent the Gaussian target pdf, the gray contour lines represent the Gaussian clutter pdf, and the straight lines represent the sensor field of view.

A schematic representing the observational geometry for the second scenario considered is given in Figure 5. The continuous state target pdf is generated by taking the continuous state target pdf from the first scenario and propagating it forward for 15 hours using the AEGIS algorithm of Reference [18]. The clutter model in this case is represented by a uniform distribution within the field-of-view of the sensor. The information gain as a function of the probability of detection is shown in Figure 6 for several values of the prior probability of object existence. In contrast to the first scenario, the information gain for different values of  $p$  do not intersect. In Figure 7, the information gain is shown as a function of prior probability of object existence for several values of the measurement noise,  $\sigma$ . Here, it is seen that lower measurement noise leads to higher information gain across the range of  $p$ . Similar to the first scenario, an inflection point is observed but with much

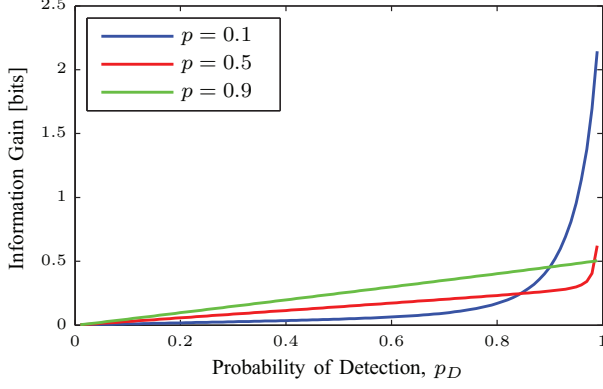


Fig. 2. Information gain as a function of probability of detection, with the measurement noise standard deviation taken to be  $\sigma = 1$  [km], and the probability of false alarm taken to be  $p_F = 0.6$ .

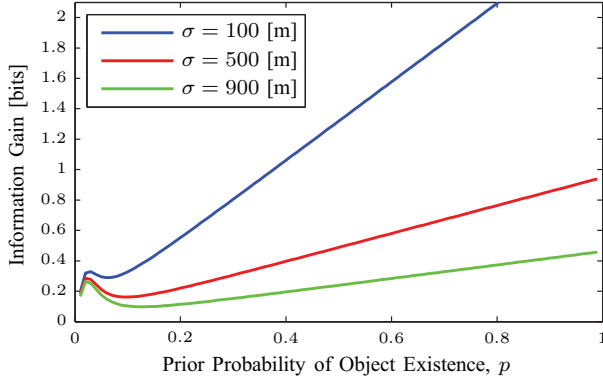


Fig. 3. Information gain as a function of prior probability of object existence, with the probability of detection taken to be  $p_D = 0.7$ , and the probability of false alarm taken to be  $p_F = 0.6$ .

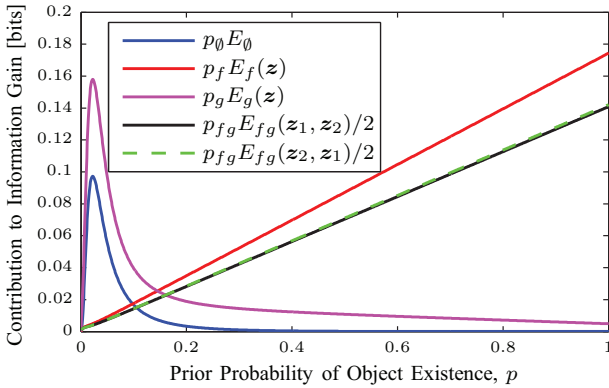


Fig. 4. Contribution of the terms in Eq. (20) to the information gain as a function of prior probability of object existence, with the probability of detection taken to be  $p_D = 0.7$ , and the probability of false alarm taken to be  $p_F = 0.6$ .

less prominence in this case. As before, the appearance of the inflection is due to the trade-off between the dominance of no return and clutter return information gain for small  $p$  and target return and two returns for large  $p$ .

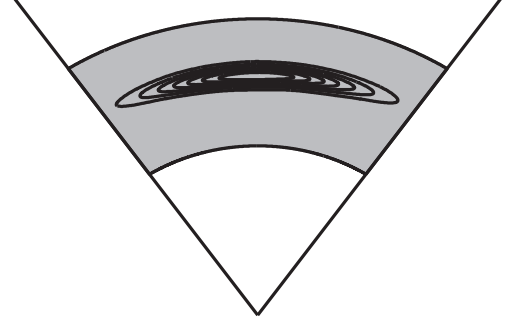


Fig. 5. Schematic of the Gaussian mixture target/uniform clutter model. The black contour lines represent the Gaussian mixture target pdf, the solid gray region represents the Gaussian clutter pdf, and the straight lines represent the sensor field of view.

## VI. CONCLUSIONS

A method for determining the information gain in multi-target tracking problems has been developed and applied for a simplified 0-1 target tracking problem. A closed-form solution for the Cauchy-Schwarz information divergence in the 0-1 target tracking problem was obtained under the assumption that the continuous state pdf is represented by a Gaussian mixture distribution and that the measurement pdf is represented by a Gaussian distribution. The information gain was applied to two scenarios in space object tracking with differing models of the continuous state distribution and the clutter pdf model. In all cases, it was shown that lower measurement noise leads to higher information gain. Similarly, it was shown that higher probability of detection leads to higher information gain. Finally, it was found that the prior probability of object

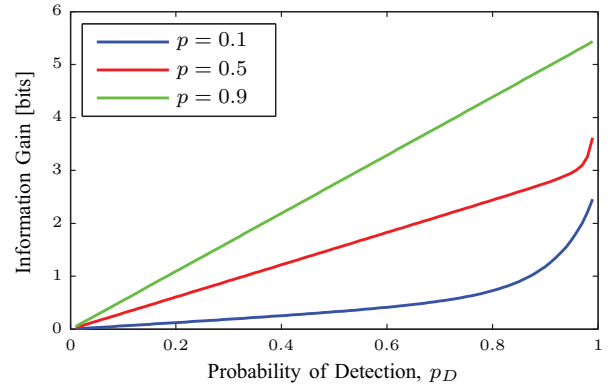


Fig. 6. Information gain as a function of probability of detection, with the measurement noise standard deviation taken to be  $\sigma = 1$  [km], and the probability of false alarm taken to be  $p_F = 0.6$ .

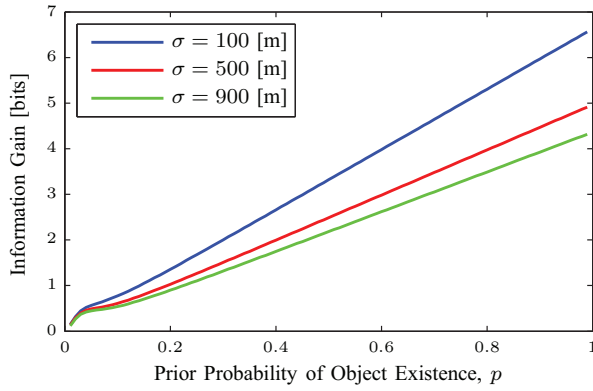


Fig. 7. Information gain as a function of prior probability of object existence, with the probability of detection taken to be  $p_D = 0.7$ , and the probability of false alarm taken to be  $p_F = 0.6$ .

existence has the most complex relationship to information gain. In some cases, a lower prior probability coupled with a high probability of detection leads to significantly more information gain, but this is not always the case as the information gain is highly situationally dependent, i.e. highly dependent on the observation geometry, observation quality, and prior continuous state distribution.

#### REFERENCES

- [1] J.-C. Liou and N. L. Johnson, "Risks in space from orbiting debris," *Science*, vol. 311, no. 5759, pp. 340–341, January 2006.
- [2] R. S. Erwin, P. Albuquerque, S. K. Jayaweera, and I. Hussein, "Dynamic sensor tasking for space situational awareness," in *Proceedings of the American Control Conference*, 2010, pp. 1153–1158.
- [3] P. S. Williams, D. B. Spencer, and R. S. Erwin, "Coupling of nonlinear estimation and dynamic sensor tasking applied to space situational awareness." Girdwood, Alaska: AAS/AIAA Astrodynamics Specialist Conference, July-August 2011.
- [4] C. Kreucher, K. Kastella, and A. O. H. III, "Information based sensor management for multitarget tracking," in *Signal and Data Processing of Small Targets*, ser. Proceedings of the SPIE, vol. 5204, 2003, pp. 480–489.
- [5] J. Aughenbaugh and B. L. Cour, "Sensor management for particle filter tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 1, pp. 503–523, January 2011.
- [6] K. J. DeMars and M. K. Jah, "Evaluation of the information content of observations with application to sensor management for orbit determination." Girdwood, Alaska: AAS/AIAA Astrodynamics Specialist Conference, July-August 2011.
- [7] I. R. Goodwin, R. P. S. Mahler, and H. T. Nguyen, *Mathematics of Data Fusion*. Kluwer Academic Publishers, 1997.
- [8] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artec House, 2007.
- [9] A. Cichocki and S. Amari, "Families of alpha- beta- and gamma-divergences: Flexible and robust measures of similarities," *Entropy*, vol. 12, pp. 1532–1568, 2010.
- [10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.
- [11] M. Uney, D. E. Clark, and S. J. Julier, "Information measures in distributed multitarget tracking," in *Proceedings of the 14th International Conference of Information Fusion*, 2011, pp. 1–8.
- [12] J. Principe, *Information Theoretic Learning: Rényi's Entropy and Kernel Perspectives*. Springer, 2010.
- [13] A. Rényi, "On measures of entropy and information," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, ser. Contributions to the Theory of Statistics, J. Neyman, Ed., vol. 1. University of California Press, June-July 1961, pp. 547–561.

- [14] R. Jenssen, "An information theoretic approach to machine learning," Ph.D. dissertation, University of Tromsø, 2005.
- [15] I. I. Hussein, K. J. DeMars, C. Früh, R. S. Erwin, and M. K. Jah, "An AEGIS-FISST integrated detection and tracking approach to space situational awareness," in *Proceedings of the 15th International Conference of Information Fusion*, 2012.
- [16] J. L. Williams and P. S. Maybeck, "Cost-function-based Gaussian mixture reduction for target tracking," in *Proceedings of the Sixth International Conference of Information Fusion*, 2003, pp. 1047–1054.
- [17] B. T. Vo, "Random finite sets in multi-object filtering," Ph.D. dissertation, The University of Western Australia, 2008.
- [18] K. J. DeMars, R. H. Bishop, and M. K. Jah, "A splitting Gaussian mixture method for the propagation of uncertainty in orbital mechanics." New Orleans, Louisiana: 21st AAS/AIAA Space Flight Mechanics Meeting, February 2011.

#### APPENDIX

For completeness, we summarize the 0-1 equations derived from FISST. The development of the following equations is treated in Reference [15]. The equations are presented for each of the possible measurement outcomes: 1) no return, 2) a single return, and 3) two returns. For each outcome, the first two equations represent the prior multi-target density function for the cases that there is no target and that there is a target, respectively. The second two equations represent the multi-target likelihood function, again when there is no target and when there is a target, respectively.

Case:  $Z_{k+1} = \emptyset$

$$\begin{aligned} f_{k+1|k}(X = \emptyset | Z^{(k)}) &= (1 - p) \\ f_{k+1|k}(X = \{\mathbf{x}\} | Z^{(k)}) &= p f_{k+1|k}(\mathbf{x} | Z^{(k)}) \\ f_{k+1}(Z_{k+1} = \emptyset | X = \emptyset) &= (1 - p_F) \\ f_{k+1}(Z_{k+1} = \emptyset | X = \{\mathbf{x}\}) &= (1 - p_F)(1 - p_D) \end{aligned}$$

Case:  $Z_{k+1} = \{\mathbf{z}\}$

$$\begin{aligned} f_{k+1|k}(X = \emptyset | Z^{(k)}) &= (1 - p) \\ f_{k+1|k}(X = \{\mathbf{x}\} | Z^{(k)}) &= p f_{k+1|k}(\mathbf{x} | Z^{(k)}) \\ f_{k+1}(Z_{k+1} = \{\mathbf{z}\} | X = \emptyset) &= p_F g(\mathbf{z}) \\ f_{k+1}(Z_{k+1} = \{\mathbf{z}\} | X = \{\mathbf{x}\}) &= p_F(1 - p_D)g(\mathbf{z}) \\ &\quad + p_D(1 - p_F)f_{k+1}(\mathbf{z} | \mathbf{x}) \end{aligned}$$

Case:  $Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\}$

$$\begin{aligned} f_{k+1|k}(X = \emptyset | Z^{(k)}) &= (1 - p) \\ f_{k+1|k}(X = \{\mathbf{x}\} | Z^{(k)}) &= p f_{k+1|k}(\mathbf{x} | Z^{(k)}) \\ f_{k+1}(Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\} | X = \emptyset) &= 0 \\ f_{k+1}(Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\} | X = \{\mathbf{x}\}) &= \\ &\quad p_F p_D (g(\mathbf{z}_1) f_{k+1}(\mathbf{z}_2 | \mathbf{x}) + g(\mathbf{z}_2) f_{k+1}(\mathbf{z}_1 | \mathbf{x})) \end{aligned}$$

In addition to the prior density and likelihood relationships, the multi-target Bayes factors are given for the no return, single return, and two return measurement outcomes, respectively, by

$$\begin{aligned} f(Z_{k+1} = \emptyset) &= (1 - p_F)(1 - p p_D) \\ f(Z_{k+1} = \{\mathbf{z}\}) &= p_F(1 - p p_D)g(\mathbf{z}) \\ &\quad + p p_D(1 - p_F)f_{k+1}(\mathbf{z}) \\ f(Z_{k+1} = \{\mathbf{z}_1, \mathbf{z}_2\}) &= p p_F p_D [g(\mathbf{z}_1) f_{k+1}(\mathbf{z}_2) \\ &\quad + g(\mathbf{z}_2) f_{k+1}(\mathbf{z}_1)] \end{aligned}$$